Some Operator and Trace Function Convexity Theorems

Eric A. Carlen[1], Rupert L. Frank[2] and Elliott H. Lieb[3]

1. Department of Mathematics, Hill Center,
Rutgers University, 110 Frelinghuysen Road, Piscataway NJ 08854-8019
2. Department of Mathematics,
Caltech, Pasadena, CA 91125
3. Departments of Mathematics and Physics, Jadwin Hall,
Princeton University, Washington Road, Princeton, NJ 08544

July 12, 2015

Abstract

We consider trace functions $(A, B) \mapsto \mathrm{Tr}[(A^{q/2}B^pA^{q/2})^s]$ where A and B are positive $n \times n$ matrices and ask when these functions are convex or concave. We also consider operator convexity/concavity of $A^{q/2}B^pA^{q/2}$ and convexity/concavity of the closely related trace functional $\mathrm{Tr}[A^{q/2}B^pA^{q/2}C^r]$. The concavity questions are completely resolved, thereby settling cases left open by Hiai; the convexity questions are settled in many cases. As a consequence, the Audenaert–Datta Rényi entropy conjectures are proved for some cases.

Mathematics subject classification numbers: 47A63, 94A17, 15A99

Key Words: Operator Convexity, Operator Concavity, Trace inequality, Rényi Entropy

1 Introduction

Let $\mathcal{P}_n$ denote the set of $n \times n$ positive definite matrices. For $p, q, s \in \mathbb{R}$, define

$$\Phi_{p,q,s}(A, B) = \mathrm{Tr}[(A^{q/2}B^pA^{q/2})^s] \ . \tag{1.1}$$

We are mainly interested in the convexity or concavity of the map $(A, B) \mapsto \Phi_{p,q,s}(A, B)$, but we are also interested in the *operator* convexity/concavity of $A^{q/2}B^pA^{q/2}$. When any of p, q or s is zero, the question of convexity is trivial, and we exclude these cases.

[1]Work partially supported by U.S. National Science Foundation grant DMS-1201354.
[2]Work partially supported by U.S. National Science Foundation grants PHY-1347399 and DMS-1363432.
[3]Work partially supported by U.S. National Science Foundation grant PHY-1265118.

Given any $n \times n$ matrix K, and with p, q, s as above, define

$$\Psi_{K,p,q,s}(A,B) = \mathrm{Tr}[(A^{q/2}K^*B^pKA^{q/2})^s] \ , \tag{1.2}$$

and note that

$$\Phi_{p,q,s}(A,B) = \Psi_{\mathbb{1},p,q,s}(A,B) \ . \tag{1.3}$$

The main question to be addressed here is this: *For which non-zero values of p, q and s is $\Psi_{K,p,q,s}(A,B)$ jointly convex or jointly concave on $\mathcal{P}_n \times \mathcal{P}_n$ for all n and all K?*

We begin with several simple reductions. Since invertible K are dense, it suffices to consider all invertible operators K. Then, for K invertible,

$$\Psi_{K,p,q,s}(A,B) = \Psi_{(K^*)^{-1},-p,-q,-s}(A,B) \ ,$$

and therefore it is no loss of generality to assume that $s > 0$. We always make this assumption in what follows.

Next, the convexity/concavity properties of $\Psi_{K,p,q,s}(A,B)$ are a consequence of those of $\Phi_{p,q,s}(A,B)$, and hence it suffices to study the special case $K = \mathbb{1}$. In fact, more is true as stated in the following Lemma 1.1. These equivalences may be useful in other contexts. (For $s = 1$ the equivalence of (1) and (4) is in [11] and the equivalence of (1) and (3) is in [4]; the arguments in those papers extend to all s, but we repeat them here for completeness.)

1.1 LEMMA (Equivalent formulations). *The following statements are equivalent for fixed p, q, s.*

(1) *The map $(A,B) \mapsto \Psi_{K,p,q,s}(A,B)$ is convex for all K and all n.*

(2) *The map $(A,B) \mapsto \Psi_{K,p,q,s}(A,B)$ is convex for all unitary K and all n.*

(3) *The map $(A,B) \mapsto \Psi_{\mathbb{1},p,q,s}(A,B) = \Phi_{p,q,s}(A,B)$ is convex for all n.*

(4) *The map $A \mapsto \Psi_{K,p,q,s}(A,A)$ is convex for all K and all n.*

(5) *The map $A \mapsto \Psi_{K,p,q,s}(A,A)$ is convex for all unitary K and all n.*

The same is true if convex is replaced by concave in all statements.

Proof. Trivially, (1) implies the other four items.

When K is unitary, $K^*A^qK = (K^*AK)^q$, and hence (3) implies (2) (even for each fixed n). By taking $K = \mathbb{1}$, (2) implies (3) (again for each fixed n).

Next we show that (2) implies (1), whence (1), (2) and (3) are equivalent. We may suppose, without loss of generality that K is a contraction. Let $K = W|K|$ be its polar decomposition. Then

$$\mathcal{U} = \begin{bmatrix} K & W\sqrt{\mathbb{1} - |K|^2} \\ -W\sqrt{\mathbb{1} - |K|^2} & K \end{bmatrix}$$

is unitary. We consider the case $q < 0$ first. For arbitrary $t > 0$, let

$$\mathcal{A}_t = \begin{bmatrix} A & 0 \\ 0 & t\mathbb{1} \end{bmatrix} \ , \ \mathcal{B} = \begin{bmatrix} B & 0 \\ 0 & 0 \end{bmatrix} .$$

Then

$$\begin{bmatrix} A^{q/2}K^*B^pKA^{q/2} & 0 \\ 0 & 0 \end{bmatrix} = \lim_{t\to\infty} \mathcal{A}_t^{q/2}\mathcal{U}^*\mathcal{B}^p\mathcal{U}\mathcal{A}_t^{q/2} .$$

Thus, recalling that we always assume $s > 0$,

$$\mathrm{Tr}[(A^{q/2}K^*B^pKA^{q/2})^s] = \lim_{t\to\infty} \Psi_{\mathcal{U},p,q,s}(\mathcal{A}_t, \mathcal{B}) .$$

Thus, (2) with $2n$ implies (1) with n. The case $q > 0$ is treated analogously, letting $t \to 0$.

Trivially, (4) implies (5). To show that (5) (with $2n$) implies (3) (with n), thereby completing the loop, replace A in (5) by $\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$, and replace K by the unitary $\begin{bmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{bmatrix}$. □

2 Known results and our extension of them

Hiai has proved in [8] that if p, q are both non-zero, and $s > 0$, and $\Phi_{p,q,s}$ is jointly *convex* in A and B, then, *necessarily*, one of the following conditions holds:

(1.) $1 \le p \le 2$ and $-1 \le q < 0$ and $s \ge 1/(p+q)$, or the same with p and q interchanged.

(2.) $-1 \le p, q < 0$ and $s > 0$.

In the special case $s = 1$, condition *(1.)* was proved to be sufficient in [1, Corollary 6.3], and condition *(2.)* was proved to be sufficient in [11, Theorem 8]; see also [3] for $s = 1$ and one of p, q negative. Hiai [8] has also proved that $\Phi_{p,q,s}$ is jointly convex in case $-1 \le p, q < 0$ and $1/2 \le s \le -1/(p+q)$.[1] Our main focus is on *(1.)*. The joint convexity in this case is known [7] when $s = 1/(p+q)$, $p = 1$ and $-1 \le q < 0$, and of course, with p and q interchanged.

Concerning *concavity,* Hiai has shown [8] that if p, q are both non-zero, and $s > 0$, and $\Phi_{p,q,s}$ is jointly concave in A and B, then, *necessarily*, the following condition holds:

(3.) $0 < p, q \le 1$ and $0 < s \le 1/(p+q)$.

In the special case $s = 1$, this condition was proved to be sufficient in [11, Theorem 1]; Hiai [8] showed sufficiency for $1/2 \le s \le 1/(p+q)$.

Our contribution to the subject is to fill in parts of the table of sufficient/necessary conditions in the following manner. We were motivated in this endeavor by a recent paper of Audenaert and Datta [2], (and Datta's Warwick lecture on it) and we prove some of their conjectures.

All the results mentioned above refer to trace inequalities. There are some *operator* convexity/concavity inequalities to be considered as well, and we will present some in the following.

[1] After this work was submitted, Hiai posted the preprint arXiv:1507.00853 in which he extended our method to prove joint convexity under condition (2.).

As far as convexity of $\Phi_{p,q,s}$ is concerned we can summarize our results as follows. We are concerned with the region $p \in [1,2]$, $q \in [-1,0)$ and $s \geq 1/(p+q)$. (Clearly, s cannot be smaller than $1/(p+q)$ by homogeneity.) We prove joint convexity for $s \geq \min\left\{\frac{1}{p-1}\ ,\ \frac{1}{1+q}\right\}$ (Thm. 4.1). Moreover, we prove joint convexity for $p = 1$ and $p = 2$ in the optimal range $s \geq 1/(p+q)$ (Thm. 4.2).

For $p \in (1,2)$, $q \in [-1,0)$, the missing regions, where we believe joint convexity also holds, is $1/(p+q) \leq s < 1$ and $1 < s < \min\left\{\frac{1}{p-1}\ ,\ \frac{1}{1+q}\right\}$. (Ando's theorem [1] covers the cases $1/(p+q) \leq s = 1$.)

On the other hand, our results completely close the gap between necessary and sufficient conditions for concavity to hold. The trace function $\Phi_{p,q,s}$ is jointly concave if and only if $0 < p, q \leq 1$ and $0 \leq s \leq 1/(p+q)$ (Thm. 4.4). This completes Hiai's results discussed above.

As for joint *operator* convexity, we prove it for $(A,B) \mapsto BA^qB$ if $-1 \leq q < 0$, and show that it does *not hold* for $(A,B) \mapsto B^{p/2}A^qB^{p/2}$ for any $p < 2$ (Thm. 3.2). (Note that it cannot hold for $p > 2$ since $B \mapsto B^p$ is not operator convex when $p > 2$.)

3 Joint operator convexity

We investigate operator convexity and concavity of certain functions on $\mathcal{P}_n \times \mathcal{P}_n$. It is well known [10, 12] that

$$(A,B) \mapsto AB^{-1}A \tag{3.1}$$

is jointly convex. In the scalar case ($n = 1$), $f(a,b) = a^qb^p$ is jointly convex on $(0,\infty) \times (0,\infty)$ if and only if $p \geq 1$, $q \leq 0$ and $p+q \geq 1$, or $q \geq 1$, $p \leq 0$ and $p+q \geq 1$, or $p, q \leq 0$. It is jointly concave if and only if $0 \leq p, q \leq 1$ and $p+q \leq 1$. It is natural to ask for which powers p and q

$$(A,B) \mapsto A^{q/2}B^pA^{q/2} \tag{3.2}$$

is jointly operator convex or concave.

This question is closely related to the question: For which values of p, q, r is

$$(A,B,C) \mapsto \mathrm{Tr} A^{q/2}B^pA^{q/2}C^r \tag{3.3}$$

jointly convex or concave in the positive operators A, B, C?

3.1 LEMMA. *When the function in (3.3) is convex (or concave) for some choice of p, q and r all non-zero, then the function in (3.2) is operator convex (or concave) for the same p and q.*

Proof. When r is positve, simply take C to be any rank-one projection. When r is negative, let P be any rank-one projection, $t > 0$. Take C to be $P + tP^\perp$, so that $C^r = P + t^rP^\perp$ and let t tend to ∞. □

Thus, the operator convexity/concavity of the operator-valued function in (3.2) is a consequence of the seemingly weaker tracial convexity/concavity of (3.3). In short, (3.3) is *stronger* than (3.2) for the same values of p, q. The value of r is irrelevant as long as it is not zero, and the implication does not even require convexity/concavity in C, only joint convexity/concavity in A and B.

When $p, r < 0$, and $-1 \leq p + r < 0$, then the map $(A, B, C) \mapsto \mathrm{Tr} A B^p A^* C^r$ is jointly convex for B, C positive and A arbitrary. This was proved in [11, Corollary 2.1]. (This *triple convexity theorem* is deeper than the double convexity theorem [11, Theorem 8] referred to in the previous section because it uses [11, Theorem 2] in an essential way.) By restricting ourselves to A positive and taking $q = 2$ this function of A, B, C reduces to (3.3).

By Lemma 3.1, the function (3.2) is jointly convex when $q = 2$ and $-1 \leq p < 0$. Our main result in this section is that there are no other cases in which this operator-valued function is either convex or concave!

3.2 THEOREM. *Let $p, q \in \mathbb{R} \setminus \{0\}$ and consider the map*

$$(A, B) \mapsto A^{q/2} B^p A^{q/2} \tag{3.4}$$

from $\mathcal{P}_n \times \mathcal{P}_n$ to $\mathcal{P}_n$ for some fixed $n \geq 2$.

(1.) The map (3.4) *is jointly operator convex if and only if $q = 2$ and $-1 \leq p < 0$.*

(2.) The map (3.4) *is* not *jointly operator concave.*

3.3 COROLLARY. *Let $p, q \in \mathbb{R} \setminus \{0\}$. The function $(A, B, C) \mapsto \mathrm{Tr} A^{q/2} B^p A^{q/2} C^r$ is never concave, and it is convex if and only if $q = 2$, $p, r < 0$ and $-1 \leq p + r < 0$.*

Proof. By Lemma 3.1, any triple convexity/concavity would imply the corresponding operator convexity/concavity, which is ruled out by the previous Theorem 3.2, except when $q = 2$, $p, r < 0$ and $-1 \leq p + r < 0$. In this case convexity is provided by [11, Corollary 2.1]. □

Our counterexamples to operator convexity and concavity given in Theorem 3.2 will be based on the following lemma.

3.4 LEMMA. *Let $r \in (-\infty, 0) \cup (0, 1)$, let $Y \geq 0$ be rank one and $n \geq 2$. Then the map $X \mapsto X^r Y X^r$ from $\mathcal{P}_n$ to $\mathcal{P}_n$ is* not *operator convex.*

Proof of Lemma 3.4. First assume that $r \in (0, 1/2)$. Then for any nontrivial $Y \geq 0$ (not necessarily rank one) the map $X \mapsto X^r Y X^r$ from $\mathcal{P}_n$ to $\mathcal{P}_n$ is not operator convex. This follows simply from the fact that the map $x \mapsto x^{2r} Y$ from $(0, \infty)$ to $\mathcal{P}_n$ is not operator convex for $0 < r < 1/2$. It is, in fact, strictly concave in this region.

Now let $r \in (-\infty, 0)$. (The proof actually also works for $r \in (0, 1/2)$, which is hardly surprising in light of the concavity mentioned above.) Clearly, we may assume $n = 2$. Let $Y = |v\rangle\langle v|$. If the convexity were true, then for all $X_1, X_2 \in \mathcal{P}_2$, with $X = (X_1 + X_2)/2$, we would have

$$X^r |v\rangle\langle v| X^r \leq \tfrac{1}{2} X_1^r |v\rangle\langle v| X_1^r + \tfrac{1}{2} X_2^r |v\rangle\langle v| X_2^r \ . \tag{3.5}$$

Without loss of generality, let $|v\rangle = (1,\,1)$. If we take $X_1 = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$ and $X_2 = t\begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}$, with $t > 0$, then (3.5) becomes

$$\begin{bmatrix} (1+t)^{2r} & (1+t)^r(1+2t)^r \\ (1+t)^r(1+2t)^r & (1+2t)^{2r} \end{bmatrix} \\ \leq 2^{2r-1}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + t^{2r}2^{2r-1}\begin{bmatrix} 1 & 2^r \\ 2^r & 2^{2r} \end{bmatrix} . \quad (3.6)$$

The vector $|w\rangle = (2^r, -1)$ is in the null space of the second matrix on the right in (3.6), and taking the trace of both sides against $|w\rangle\langle w|$ yields

$$\left\langle w, \begin{bmatrix} (1+t)^{2r} & (1+t)^r(1+2t)^r \\ (1+t)^r(1+2t)^r & (1+2t)^{2r} \end{bmatrix} w \right\rangle \leq 2^{2r-1}(2^r-1)^2 ,$$

which, in the limit $t \to 0$, becomes $(2^r-1)^2 \leq 2^{2r-1}(2^r-1)^2$, so that for $r \neq 0$, we would have $1 \leq 2^{2r-1}$. This is false for all $r < 1/2$, which shows that (3.5) leads to a contradiction for nonzero $r \in (-\infty, 0) \cup (0, 1/2)$.

Our proof for $1/2 \leq r < 1$ is different; this proof actually works in the range $0 < r < 1$. Let $|v\rangle$ be a unit vector in $\mathbb{C}^n$. Then we will show that there is another vector $|w\rangle$ in $\mathbb{C}^n$ such that

$$X \mapsto |\langle w|X^r|v\rangle|^2$$

is not convex. Again, we may assume that $n = 2$ and that $|v\rangle = (0, 1)$. Take

$$X_1 = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} \quad \text{and} \quad X_2 = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} .$$

Let $|w\rangle = (1, -1)$, so that $X_1^r|w\rangle = 0$ and $X_2^r|v\rangle = 0$. Evidently,

$$\tfrac{1}{2}|\langle w|X_1^r|v\rangle|^2 + \tfrac{1}{2}|\langle w|X_2^r|v\rangle|^2 = 0 .$$

However, the eigenvalues of $X = \frac{1}{2}(X_1 + X_2)$ are easily computed to be $\lambda_\pm = (3 \pm \sqrt{5})/2$, and then a further simple computation yields

$$\langle w|X^r|v\rangle = \frac{1}{\sqrt{5}}(\lambda_+^{r-1} - \lambda_-^{r-1}) ,$$

and this is strictly negative for all $0 < r < 1$. □

Proof of Theorem 3.2. As explained above, the convexity assertion in *(1.)* is a consequence of [11, Corollary 2.1]. Our goal now is to prove that there are no other cases of convexity or concavity.

A number of exponents can be excluded by considering the scalar case. Moreover, since $X \mapsto X^r$ is operator convex on $\mathcal{P}_n$ if and only if $r \in [-1, 0] \cup [1, 2]$, and is operator concave on $\mathcal{P}_n$ if and only if $r \in [0, 1]$, the only cases in which convexity cannot be immediately ruled out are $p \in [1, 2]$, $q \in [-1, 0]$ and $p + q \geq 1$ (or the same with p and q interchanged). Likewise, the only cases of in which concavity cannot be immediately ruled out are $p, q \in [0,1]$, $p+q \leq 1$.

For part *(1.)* it remains for us to show that (3.4) is not jointly operator convex in the following three cases,

(a) $p \in [-1,0)$, $q \in [1,2)$ and $p+q \geq 1$.

(b) $p \in [1,2]$, $q \in [-1,0)$ and $p+q \geq 1$.

(c) $p \in (-1,0)$ and $p+q \geq -1$.

Let us prove failure of convexity in case (a). Let $|v\rangle$ be any unit vector in $\mathbb{C}^n$. Let P be the orthogonal projection onto the span of v, and let $P^\perp$ denote the complementary projection. Fix $t > 0$, and define $B_t = P + tP^\perp$. Then $B_t^p = P + t^p P^\perp$. If convexity would hold, then for any $|w\rangle$ the map $A \mapsto \langle w|A^{q/2} B_t^p A^{q/2}|w\rangle$ would be convex. Since $\lim_{t\to\infty} B_t^p = |v\rangle\langle v|$, and since limits of convex functions are convex, it would follow that $A \mapsto |\langle v|A^{q/2}|w\rangle|^2$ would be convex on $\mathcal{P}_n$ for any $|w\rangle$. This contradicts Lemma 3.4 with $r = q/2 \in [1/2, 1)$. The proof for (c) is almost exactly the same, except one uses Lemma 3.4 with $r = q/2 < 0$.

The proof in case (b) is similar. Again, we let $|v\rangle$ be a unit vector in $\mathbb{C}^n$ and set $B = |v\rangle\langle v|$. Then $B^p = |v\rangle\langle v|$ and, if convexity would hold, then for any $|w\rangle$ the map $A \mapsto |\langle v|A^{q/2}|w\rangle|^2$ would be convex on $\mathcal{P}_n$. This contradicts Lemma 3.4 with $r = q/2 \in [-1/2, 0)$.

Finally, we prove *(2.)*, the failure of concavity. According to the discussion above, it remains for us to show that (3.4) is not jointly operator concave for $p, q \in (0,1]$ and $p+q \leq 1$. Suppose $(A,B) \mapsto A^{q/2}B^pA^{q/2}$ were concave for some p, q in this range. Then for all non-negative A and B we would have

$$\begin{aligned} \tfrac{1}{2}A^{q/2}B^pA^{q/2} + \tfrac{1}{2}B^{q/2}A^pB^{q/2} &\leq \left(\frac{A+B}{2}\right)^{q/2}\left(\frac{B+A}{2}\right)^{p}\left(\frac{A+B}{2}\right)^{q/2} \\ &= 2^{-p-q}(A+B)^{p+q} \ . \end{aligned}$$

Suppose that A has a non-trivial null space (here we use the assumption $n \geq 2$), and $|v\rangle$ is a unit vector with $A|v\rangle = 0$. By Jensen's inequality, since $p+q \leq 1$,

$$\langle v|(A+B)^{p+q}|v\rangle \leq \langle v|(A+B)|v\rangle^{p+q} = \langle v|B|v\rangle^{p+q} \ .$$

Thus we would have

$$\langle v|B^{q/2}A^pB^{q/2}|v\rangle \leq 2^{1-p-q}\langle v|B|v\rangle^{p+q} \ .$$

The left side is homogeneous of degree q in B, while the right side is homogeneous of degree $p+q$, and hence the inequality cannot be generally valid. (The positivity of the powers is essential here; the argument of course cannot be adapted to yield a counterexample to the convexity proved in the first part of the theorem.) $\square$

3.5 Remark. There is another way to prove the convexity in (3.4) for $q = 2$ and $-1 \leq p < 0$. For $p = -1$ one can use the Schwarz type inequality in [12, 10]. (This inequality, however, is actually weaker than the triple convexity inequality [11, Corollary 2.1] that we used in the proof of Theorem 3.2.) For $-1 < p < 0$ one can use the integral representation $B^p = C_p \int_0^\infty (B+t)^{-1} t^p \mathrm{d}t$ with $C_p > 0$ to reduce matters to the case $p = -1$. Indeed, one can replace B^p by any Herglotz function $\int_{t\geq 0} (B+t)^{-1} \mathrm{d}\mu(t)$ with $\mu > 0$.

4 Convexity of $\Phi_{p,q,s}(A, B)$

In this section we prove, among other things, two cases of a conjecture of Audenaert and Datta [2]. Much of our analysis is based on the formulas

$$\mathrm{Tr}[X^s] = s \sup_{Z\geq 0} \left\{ \mathrm{Tr}[XZ^{1-1/s}] + \left(\tfrac{1}{s} - 1\right) \mathrm{Tr}[Z] \right\} \qquad \text{if } s > 1 \tag{4.1}$$

and

$$\mathrm{Tr}[X^s] = s \inf_{Z>0} \left\{ \mathrm{Tr}[XZ^{1-1/s}] + \left(\tfrac{1}{s} - 1\right) \mathrm{Tr}[Z] \right\} \qquad \text{if } 0 < s < 1\,; \tag{4.2}$$

see [4, Lemma 2.2]. These formulas have already played an important role in our previous works [4] and [7].

4.1 THEOREM. *When $p \in [1, 2]$, $q \in [-1, 0)$, $\Phi_{p,q,s}(A, B)$ is jointly convex for all*

$$s \geq \min\left\{ \frac{1}{p-1} \ , \ \frac{1}{1+q} \right\} \ .$$

Here we set $\frac{1}{p-1} = +\infty$ for $p = 1$ and $\frac{1}{1+q} = +\infty$ for $q = -1$. Thus, the theorem implies that, in particular, for $p = 1$, $\Phi_{1,q,s}(A, B)$ is jointly convex in the optimal range $q \in [-1, 0)$ and $s \geq \frac{1}{1+q}$. An optimal result for $p = 2$ will be proved in Theorem 4.2. As discussed in Section 2, for $p \in (1, 2)$, $q \in [-1, 0)$, the region where convexity is not settled is $1/(p + q) \leq s < 1$ and $1 < s < \min\left\{ \frac{1}{p-1} \ , \ \frac{1}{1+q} \right\}$.

Proof. First, we prove convexity if $s \geq 1/(1 + q)$. Since this implies $s > 1$, we have by (4.1),

$$\Phi_{p,q,s}(A, B) = s \sup_{Z\geq 0} \left\{ \mathrm{Tr}[A^{q/2} B^p A^{q/2} Z^{1-1/s}] + \left(\tfrac{1}{s} - 1\right) \mathrm{Tr}[Z] \right\} \ .$$

Now define $D^2 = A^{q/2} Z^{(s-1)/s} A^{q/2}$ and note that $Z = (A^{-q/2} D^2 A^{-q/2})^{s/(s-1)}$ to write

$$\Phi_{p,q,s}(A, B) = s \sup_{D\geq 0} \left\{ \mathrm{Tr}[DB^p D] + \left(\tfrac{1}{s} - 1\right) \mathrm{Tr}[(DA^{-q} D)^{s/(s-1)}] \right\} \ . \tag{4.3}$$

For $1 \leq p \leq 2$, the map $B \mapsto B^p$ is operator convex and therefore $B \mapsto \mathrm{Tr}[DB^p D]$ is convex. Moreover, by Hiai's extension of Epstein's Theorem [8, Thm. 4.1] the map $A \mapsto \mathrm{Tr}[(DA^{-q} D)^{s/(s-1)}]$ is concave as long as $s/(s-1) \leq -1/q$, which is the same as $s \geq 1/(1+q)$. Thus, (4.3) represents $\Phi_{p,q,s}(A, B)$ as a supremum of jointly convex functions and so $\Phi_{p,q,s}(A, B)$ is jointly convex for $s \geq 1/(1+q)$. This proves the first part of the theorem.

We now prove convexity if $s \geq 1/(p-1)$. Let us first consider the case $p = 2$ and $s = 1$, where $\Phi_{2,q,1}(A,B) = \mathrm{Tr}[A^{q/2}B^2A^{q/2}] = \mathrm{Tr}[BA^qB]$. For $-1 \leq q < 0$, the map $(A,B) \mapsto BA^qB$ is operator convex by Theorem 3.2 and therefore $(A,B) \mapsto \mathrm{Tr}[BA^qB]$ is convex, as claimed. We now assume that $s > 1$ (and still $s \geq 1/(p-1)$). Then by (4.1), making use of $\mathrm{Tr}[(A^{q/2}B^pA^{q/2})^s] = \mathrm{Tr}[(B^{p/2}A^qB^{p/2})^s]$,

$$\Phi_{p,q,s}(A,B) = s \sup_{Z \geq 0} \left\{ \mathrm{Tr}[B^{p/2}A^qB^{p/2}Z^{1-1/s}] + \left(\tfrac{1}{s}-1\right)\mathrm{Tr}[Z] \right\} .$$

Note that

$$\mathrm{Tr}[B^{p/2}A^qB^{p/2}Z^{1-1/s}] = \mathrm{Tr}[BA^qB(B^{p/2-1}Z^{1-1/s}B^{p/2-1})] .$$

Define $D^2 = B^{p/2-1}Z^{(s-1)/s}B^{p/2-1}$, so that $Z = (B^{1-p/2}D^2B^{1-p/2})^{s/(s-1)}$. Then

$$\begin{aligned}\Phi_{p,q,s}(A,B) &= s \sup_{D \geq 0} \left\{ \mathrm{Tr}[DBA^qBD] + \left(\tfrac{1}{s}-1\right)\mathrm{Tr}[(B^{1-p/2}D^2B^{1-p/2})^{s/(s-1)}] \right\} \\ &= s \sup_{D \geq 0} \left\{ \mathrm{Tr}[DBA^qBD] + \left(\tfrac{1}{s}-1\right)\mathrm{Tr}[(DB^{2-p}D)^{s/(s-1)}] \right\} .\end{aligned} \tag{4.4}$$

Since $-1 \leq q < 0$, $(A,B) \mapsto BA^qB$ is operator convex by Theorem 3.2, so $(A,B) \mapsto \mathrm{Tr}[DBA^qBD]$ is convex. By Hiai's extension of Epstein's Theorem [8, Thm. 4.1], $B \mapsto \mathrm{Tr}[(DB^{2-p}D)^{s/(s-1)}]$ is concave as long as $s/(s-1) \leq 1/(2-p)$, which is the same as $s \geq 1/(p-1)$. Thus, (4.4) represents $\Phi_{p,q,s}(A,B)$ as a supremum of jointly convex functions and so $\Phi_{p,q,s}(A,B)$ is jointly convex for $s \geq 1/(p-1)$. This completes the proof. □

4.2 THEOREM. *When $p = 2$, $\Phi_{p,q,s}(A,B)$ is jointly convex for all $-1 \leq q < 0$ and $s \geq 1/(2+q)$.*

This result yields the optimal range of convexity for $p = 2$. It had been conjectured in [2] for $s = 1/(2+q)$.

Proof. The convexity for $s \geq 1$ follows from Theorem 4.1 and therefore we may assume that $1/(p+q) \leq s < 1$. Then, making use of $\mathrm{Tr}[(A^{q/2}B^2A^{q/2})^s] = \mathrm{Tr}[(BA^qB)^s]$,

$$\Phi_{2,q,s}(A,B) = s \inf_{Z > 0} \left\{ \mathrm{Tr}[BA^qBZ^{1-1/s}] + \left(\tfrac{1}{s}-1\right)\mathrm{Tr}[Z] \right\} . \tag{4.5}$$

The important distinction between this formula and formulas (4.3) and (4.4) is the infimum in place of the supremum. Joint convexity in A, B no longer suffices. Instead we need joint convexity in A, B, Z, with which we can apply [4, Lemma 2.3].

Note that $1 - 1/s \leq 0$. By [11, Corollary 2.1], $(A,B,Z) \mapsto \mathrm{Tr}[BA^qBZ^{1-1/s}]$ is jointly convex as long as $q + 1 - 1/s \geq -1$, which means $s \geq 1/(2+q)$. For such s, the argument of the infimum in (4.5) is jointly convex in A, B and Z. By [4, Lemma 2.3], the infimum itself is jointly convex in A and B. This proves the assertion for $1/(2+q) \leq s \leq 1$. □

4.3 Remark. In the previous proof for the range $s \geq 1$ we referred to Theorem 4.1 which, in turn, was based on Hiai's extension of Epstein's theorem. For the case relevant for Theorem 4.2, however, there is a more direct proof. Indeed, let $A_j, B_j \in \mathcal{P}_n$, $j = 1, 2$, and $\lambda \in (0, 1)$ and set $A = \lambda A_1 + (1 - \lambda)A_2$ and $B = \lambda B_1 + (1 - \lambda)B_2$. Then by Theorem 3.2 for $-1 \leq q < 0$,

$$BA^qB \leq \lambda B_1 A_1^q B_1 + (1 - \lambda)B_2 A_2^q B_2 \ .$$

For all $s \geq 0$, $X \mapsto \mathrm{Tr}[X^s]$ is monotone on $\mathcal{P}_n$. Hence, even for all $s \geq 0$,

$$\mathrm{Tr}[(BA^qB)^s] \leq \mathrm{Tr}[(\lambda B_1 A_1^q B_1 + (1 - \lambda)B_2 A_2^q B_2)^s] \ .$$

Finally, for $s \geq 1$, $X \mapsto \mathrm{Tr}[X^s]$ is convex on $\mathcal{P}_n$. Therefore,

$$\mathrm{Tr}[(\lambda B_1 A_1^q B_1 + (1-\lambda)B_2 A_2^q B_2)^s] \leq \lambda \mathrm{Tr}[(B_1 A_1^q B_1)^s] + (1-\lambda)\mathrm{Tr}[(B_1 A_1^q B_1)^s] \ .$$

This proves the convexity for $s \geq 1$ and $-1 \leq q < 0$.

The next result concerns the concavity of $\Phi_{p,q,s}(A, B)$.

4.4 THEOREM. *The trace function $\Phi_{p,q,s}(A, B)$ is jointly concave if and only if $0 \leq p, q \leq 1$ and $0 \leq s \leq 1/(p + q)$.*

Proof. The necessity of the condition is proved in [8, Prop. 5.1] and the sufficiency for $1/2 \leq s \leq 1/(p + q)$ is proved in [8, Thm. 2.1]. Our task is to prove sufficiency in the case $0 < s < 1/2$. We write, using (4.2),

$$\begin{aligned}
\Phi_{p,q,s}(A, B) &= s \inf_{X>0} \mathrm{Tr} \left\{ A^{q/2} B^p A^{q/2} X^{1-1/s} + (\tfrac{1}{s} - 1)X \right\} \\
&= s \inf_{Y>0} \mathrm{Tr} \left\{ B^p Y + (\tfrac{1}{s} - 1) \left(A^{q/2} Y^{-1} A^{q/2} \right)^{s/(1-s)} \right\} \\
&= s \inf_{Y>0} \mathrm{Tr} \left\{ B^p Y + (\tfrac{1}{s} - 1) \left(Y^{-1/2} A^q Y^{-1/2} \right)^{s/(1-s)} \right\} \ .
\end{aligned}$$

Since $0 \leq p \leq 1$, $B \mapsto B^p$ is operator concave and so $B \mapsto \mathrm{Tr}B^pY$ is concave. By the extension of Epstein's Theorem proved in [8, Theorem 4.1], $A \mapsto \mathrm{Tr}(Y^{-1/2}A^qY^{-1/2})^{s/(1-s)}$ is concave if $s/(1 - s) \leq 1/q$. This condition is satisfied since $s \leq 1/2 \leq 1/(1 + q)$. We conclude that $\Phi_{p,q,s}(A, B)$ as an infimum of concave functions is concave. □

We conclude with a corollary of Theorem 4.2. For $\rho, \sigma \in \mathcal{P}_n$ and $\alpha, z > 0$, we introduce the so-called $\alpha - z$-relative Rényi entropies

$$D_{\alpha,z}(\rho||\sigma) = \frac{1}{\alpha - 1} \ln \frac{\mathrm{Tr} \left(\sigma^{(1-\alpha)/(2z)} \rho^{\alpha/z} \sigma^{(1-\alpha)/(2z)} \right)^z}{\mathrm{Tr}\rho} \ .$$

(For $\alpha = 1$, a limit has to be taken.) These functionals appeared in [9, Sec. 3.3] and were further studied in [2], where the question was raised whether the $\alpha - z$-relative Rényi entropies are monotone under completely positive, trace preserving maps. Currently this is known for $0 < \alpha \leq 1$ and $z \geq \max\{\alpha, 1 - \alpha\}$, and for $1 \leq \alpha \leq 2$ and $z = 1$, and for $1 \leq \alpha < \infty$ and $z = \alpha$. See [2] for these cases. In this paper Audenaert and Datta conjecture that monotonicity holds for $1 \leq \alpha \leq 2$ and $\alpha/2 \leq z < \alpha$, and for $2 \leq \alpha < \infty$ and $\alpha - 1 \leq z < \alpha$. Our contribution here is to prove their conjecture for $1 < \alpha = 2z \leq 2$.

4.5 COROLLARY. *Let $\alpha = 2z \in (1,2]$ and let $\rho, \sigma \in \mathcal{P}_n$. Then for any completely positive, trace preserving map $\mathcal{E}$ on $\mathcal{P}_n$,*

$$D_{\alpha,\alpha/2}(\rho||\sigma) \geq D_{\alpha,\alpha/2}(\mathcal{E}(\rho)||\mathcal{E}(\sigma))\,.$$

Proof. By a classical argument due to Lindblad and Uhlmann, see, e.g., [5, 7], the monotonicity follows once it is shown that

$$(\rho,\sigma) \mapsto \mathrm{Tr}\left(\sigma^{(1-\alpha)/\alpha}\rho^2\sigma^{(1-\alpha)/\alpha}\right)^{\alpha/2} = \Phi_{2,2(1-\alpha)/\alpha,\alpha/2}(\sigma,\rho)$$

is jointly convex. For $\alpha \in (1,2]$ this convexity follows from Theorem 4.2. $\square$

Acknowledgements We thank Marius Lemm and Mark Wilde, as well as the anonymous referee, for useful remarks.

References

[1] T. Ando, *Concavity of certain maps on positive definite matrices and applications to Hadamard products*, Lin. Alg. and its Appl. **26**, 203–241 (1979).

[2] K. M. R. Audenaert and N. Datta, *α-z-relative Renyi entropies*, Jour. Math. Phys., **56**, 022202 (2015). arXiv:1310.7178

[3] T. N. Bekjan, *On joint convexity of trace functions*, Lin. Alg. and its Appl. **390**, 321–327 (2004).

[4] E. A. Carlen and E. H. Lieb, *A Minkowski type trace inequality and strong subadditivity of quantum entropy II: convexity and concavity*, Lett. Math. Phys. **83**, 107–126 (2008). arXiv:0710.4167

[5] E. A. Carlen, *Trace inequalities and quantum entropy: an introductory course*, in: *Entropy and the quantum*, 73–140, Contemp. Math. **529**, Amer. Math. Soc., Providence, RI, 2010.

[6] H. Epstein, *Remarks on two theorems of E. Lieb*, Commun. Math. Phys. **31**, 317–325 (1973).

[7] R. L. Frank and E. H. Lieb, *Monotonicity of a relative Renyi entropy*, Jour. Math. Phys. **54**, 122201 (2013). DOI: 10.1063/1.4838835. arXiv:1306.5358

[8] F. Hiai, *Concavity of certain matrix trace and normed functions*, Lin. Alg. and its Appl. **439**, 1568–1589 (2013). arXiv:1210.7524

[9] V. Jaksic, Y. Ogata, Y. Pautrat and C.-A. Pillet, *Entropic fluctuations in quantum statistical mechanics. An Introduction.* In: Quantum Theory from Small to Large Scales: Lecture Notes of the Les Houches Summer School: Volume **95**, August 2010, Oxford University Press, 2012.

[10] J. Kiefer, *Optimum experimental designs*, J. Roy. Statist. Soc. Ser. B **21**, 272–310 (1959).

[11] E. H. Lieb, *Convex trace functions and the Wigner-Yanase-Dyson conjecture*, Adv. in Math. **11**, 267–288 (1973).

[12] E. H. Lieb and M. B. Ruskai, *Some operator inequalities of the Schwarz type*, Adv. in Math. **12**, 269–273 (1974).